Received: 20 October 2017 | Revised: 4 May 2018 | Accepted: 8 May 2018

DOI: 10.1002/mpr.1725

INVITED REVIEW

WILEY

Prospective: Is bipolar disorder being overdiagnosed?

Tammas Kelly[1,2]

[1]Department of Psychiatry and Behavioral Sciences, George Washington University, Washington DC, USA

[2]The Depression & Bipolar Clinic of Colorado, Fort Collins, Colorado, USA

Correspondence
Tammas Kelly, Department of Psychiatry and Behavioral Sciences, George Washington University, 2120 L St NW, Suite 600, Washington DC 20037, USA.
Email: tamkelly@comcast.net

Abstract

Objectives: Many studies indicate that bipolar disorders are underdiagnosed. Yet from 2007 to 2008, a series of publications asserted that bipolar disorders were being overdiagnosed. This review examines the methods used in the studies that reported bipolar disorders were being overdiagnosed.

Methods: A literature search for studies with original data related to overdiagnosis of bipolar disorders was performed.

Results: Four studies were found indicating bipolar disorders were being overdiagnosed. The Structured Clinical Interview of the Diagnostic and Statistical Manual of Mental Disorders (SCID) was used in the diagnostic process. The studies compared the clinical diagnosis of bipolar disorder to a single SCID interview without interviewing family or reviewing old records. The studies assumed the SCID diagnosis was correct.

Conclusions: Numerous concerns were found. The SCID frequently missed diagnosis of bipolar, the definitions of bipolar disorder are so narrow and conservative that the outcomes of the studies may have been predetermined. Ultimately, the studies compared the strength of a diagnosis made by a treating psychiatrist to a SCID diagnosis collected with virtually no information from the clinician. The assumption that the SCID diagnosis is always correct and the clinician is always wrong is unsupportable. The premise that bipolar disorders are being overdiagnosed is unsupported by reasonable science.

KEYWORDS

bipolar disorders, clinicians, overdiagnosis, SCID, underdiagnosis

1 | INTRODUCTION

Methodological paradigms, research models, and statistical methods are ever evolving in the ongoing pursuit to assure the highest standards of research, nosology, and treatment. It is critical that we evaluate and adopt the tools that yield the most accurate information. However, we should not lose sight of the fundamentals of what constitutes good research. Are the study methods adequate? Are the instruments appropriate? Does the nosology conform to recognized standards? If not, is the research clear about how the nosology differs from the standards? Do the results vary significantly from what could be reasonably expected? In comparison studies, what constitutes sufficient data to make a true comparison? How much latitude should be given to speculations by the authors? This review illustrates the importance of these fundamentals.

The recognition and understanding of bipolar disorders has been growing since the introduction of lithium facilitated the differentiation of bipolar from schizophrenia. As our knowledge has increased, our ability to diagnose bipolar disorders has also improved. Some are uncomfortable with what they believe are ever-expanding boundaries and have stated, "These efforts have given way to a modern-day epidemic of bipolar disorder" (Goldberg, 2010). They fear diagnostic chaos and the difficulties that "fuzzy" boundaries pose to research (Goldberg, 2015). A fundamental question arises—namely, are the boundaries of bipolar expanding or just our understanding? If we consider Kraepelin's work, the boundaries of bipolar disorder are not

expanding. Kraepelin's original 1898 definition of manic-depression cast a wide net:

> *Manic-depressive insanity … includes on the one hand the whole domain of so-called periodic and circular insanity, and the other hand simple mania, the greater part of the morbid state termed melancholy and also a not incomprehensible number of cases of amentia (confusional or delirious insanity), lastly, we include here certain slight and slightest colorings of mood, some of them periodic, some of them continuously morbid, which on the one hand are to be regarded as the rudiment of more severe disorders, on the other hand pass over without sharp boundary into the domain of personal predisposition. (Kraepelin, 1921/1898)*

As a superintendent of a large psychiatric hospital in what is now modern-day Estonia, Kraepelin was able to make long-term observations of a large number of patients suffering from bipolar disorders unencumbered by effective treatments.

Many studies have shown that bipolar disorders are underdiagnosed (Akiskal et al., 1983; Akiskal et al., 1995; Albanese, Clodfelter Jr, Pardo, & Ghaemi, 2006; Benazzi, 2003a; Dunner, Gershon, & Goodwin, 1976; Dunner & Tay, 1993; Ghaemi, Boiman, & Goodwin, 2000; Ghaemi, Sachs, Chiou, Pandurangi, & Goodwin, 1999; Hirschfeld et al., 2003; Hirschfeld, Lewis, & Vornik, 2003; Smith et al., 2011; Zimmerman, Ruggero, Chelminski, & Young, 2008). For many years, this was the prevailing view. As the diagnosis of bipolar became more common, many psychiatrists became uneasy with the increase in diagnosed cases. A series of three influential studies, published from 2007 to 2008, asserted that overdiagnosis of bipolar disorder is a problem. Clinically, this assertion has led patients, nonpsychiatric physicians, psychologists, therapist, and many psychiatrists to discount the diagnosis of bipolar disorder.

Although other reviews have addressed the literature regarding overdiagnosis, no review has examined the methods used to assess the overdiagnosis (Ghouse, Sanches, Zunta-Soares, Swann, & Soares, 2013; Mitchell, 2012). This review examines the methods used in the studies that reported bipolar disorders were being overdiagnosed.

2 | METHODS

Google Scholar and PubMed were used to search for all relevant articles pertaining to the overdiagnosis of bipolar disorders in adults that contained unique data sets. Studies using previous data sets were excluded. Once relevant articles were identified, the citations within the identified articles were examined other studies that may be relevant using the PubMed “Related Citations” feature to identify other relevant articles.

3 | RESULTS

Three studies containing original data that assert the overdiagnosis of the bipolar disorders were found (Goldberg et al., 2008; Stewart & El-Mallakh, 2007; Zimmerman et al., 2008). Data from a fourth study were used to assert overdiagnosis of bipolar disorders, but the study itself makes no such claim (Hirschfeld, Cass, Holt, & Carlson, 2005; Zimmerman, 2010).

The Stewart and El-Mallakh (Stewart) study is a retrospective analysis of 21 consecutive patients entering a residential substance abuse program who were previously diagnosed with a bipolar disorder. The diagnoses of bipolar not otherwise specified (NOS) and cyclothymia were not considered. Despite this exclusion, it does not appear that the study sample was limited to only patients previously diagnosed with Bipolar I or II. Based on a single Structured Clinical Interview of the Diagnostic and Statistical Manual of Mental Disorders (SCID), 52.4% of patients fail to be diagnosed with Bipolar I or II and are diagnosed with substance-induced mood disorders with manic or hypomanic features (Stewart & El-Mallakh, 2007).

The Goldberg et al. (Goldberg) study is a retrospective review of 85 hospital charts of patients previously diagnosed with bipolar disorders. The diagnosis is based on a single interview, but some information may have been gathered from a preadmission interview with the patient's outpatient clinicians and a meeting with the patient's family. There is no indication that these preinterviews, conducted by social workers, systematically collected information regarding bipolar disorders or contributed to the diagnostic process. Curiously, despite the authors' repeated mention of the SCID, it does not appear that the SCID was actually used. As in the Stewart study, the final analysis only considers patients with the diagnosis of Bipolar I or II disorders. However, it does not appear that the authors made any attempt to exclude patients diagnosed as bipolar NOS or cyclothymia. Patients failed to be diagnosed with bipolar due to the following overlapping reasons: 55% insufficient criteria B symptoms, 12% failure to meet duration criteria, 36% lacked periods of abstinence to allow for diagnosis of bipolar, and 63% of patients had mood symptoms solely during a 4-week period after significant substance use. Only 33% of patients previously diagnosed with a bipolar disorder met full criteria for Bipolar I or II. The authors concluded that 67% of patients were overdiagnosed with bipolar disorder (Goldberg et al., 2008).

The Hirschfeld et al. (Hirschfeld) study makes no assertion that bipolar disorders were being overdiagnosed, but Zimmerman et al. use this study as proof of overdiagnosis (Hirschfeld et al., 2005; Zimmerman, 2010). Hirschfeld used psychiatric social workers trained in the use of the SCID to interview 180 patients by phone using a modified SCID. The patients selected were outpatients from a family medicine clinic at the University of Texas who were being treated with antidepressants. Of the 43 patients self-identifying as previously diagnosed with a bipolar disorder, 32.6% were not diagnosed with a bipolar disorder. Of the patients who did not self-identify as having a history of bipolar disorder, 21.9% were diagnosed with bipolar (Hirschfeld et al., 2005).

The Zimmerman et al. (Zimmerman) study is a retrospective analysis of 700 charts of patients who underwent a single SCID interview by bachelor's level technician or PhD psychologist. Of the 700 patients, 20.7% (N = 145) self-identified as having previously been diagnosed with a bipolar disorder. In more than half of these patients (56.6%), the diagnosis could not be confirmed. These patients were considered overdiagnosed. Of the patients not previously diagnosed with bipolar disorder, 4.7% were diagnosed as bipolar and considered

underdiagnosed. Using the Family History Research Diagnostic Criteria to gather family history, Zimmerman asserts that family histories validate their findings with the SCID. The Family History Research Diagnostic Criteria method attempts to directly assess symptoms of psychiatric illnesses based on an interview with the patient that specifically tries to solicit the presence or absence of symptoms associated with the most common mental illnesses. However, in the Zimmerman study, the collection of family history was based solely on the patient's recollection. There was no systematic information collected directly from family members (Zimmerman et al., 2008).

None of the four studies gathered systematic family history directly from family or systematically reviewed past records, and the studies did not analyze data regarding illness features that have been repeatedly associated with a bipolar diagnosis, such as age of onset of first depression, course of illness, postpartum onset, psychosis, repeated failures to respond to antidepressants, a hypomanic/manic response to an antidepressant, or a high recurrence rate of depressive episodes. This approach is difficult to understand given these methods could be considered as more objective and easily obtained. A more thorough review would have assessed all of these factors (Ghaemi, Ko, & Goodwin, 2002; Mitchell, Goodwin, Johnson, & Hirschfeld, 2008; Phelps, Angst, Katzow, & Sadler, 2008).

4 | DISCUSSION

These four studies are the only ones that purport to show that bipolar disorders are overdiagnosed. A close review reveals significant concerns that fall into four broad areas: limitations of information gathered, limitations of methods, limitations of the conservative boundaries set for the diagnosis of bipolar disorders, and assertions offered without proof.

4.1 | Limitations of information gathered

In the normal course of clinical practice, most clinicians would be reluctant to rule out bipolar based on a single interview. Further interviews would provide the opportunity to gather additional information, including old records, family observations, more history, and direct observation of symptoms.

4.1.1 | Factors that make bipolar disorder difficult to diagnose

There are many well-known factors that can interfere with the diagnostic process. Patients with bipolar disorder often lack insight into their illness or even that they have an illness. This lack of insight is comparable with that found in schizophrenia (Braw et al., 2012). Insight into Bipolar II is worse than Bipolar I (Braw et al., 2012; Látalová, 2012). State-dependent memory can also be an important factor. The DSM-IV-TR states that "Often individuals, particularly when in the midst of a Major Depressive Episode, do not recall periods of hypomania" (APA & DSM-IV., 1994). Another factor is many patients either do not understand or infer different meanings to words such as euphoria, up, elevated, expensive, overly happy, exuberant, racing thoughts, and so forth. One of the first tasks for psychiatrists and patients is to develop a common language, and this takes time.

Denial is an ever-present problem in psychiatry. Patients do not want the stigma associated with bipolar, believing that the label will adversely affect their lives. Some patients simply do not want to be bipolar. The reasons range from not wanting to be "like" a parent who suffered from a bipolar disorder to shame. It is not unusual for families to keep mental illnesses secret. Some families are simply stoic and/or do not believe mental illnesses are real or valid. These families often model suppression or denial of emotions, making it difficult to elicit an accurate history. Families often withhold mental health history altogether from their children.

Transference issues are also ever present and can interfere with the diagnostic process. Similarly, "wanting to please the interviewer" can impede the diagnostic process. Wanting to please the interviewer may be a particular problem in inpatient or residential settings where patients believe their length of stay is dependent on the interviewer. In clinical practice, these factors decrease and more history is gathered as patients develop trust in their clinician, develop a common language with their clinician, and are directly observed.

4.1.2 | The importance of past records and information from family

None of the four studies conducted systematic family interviews looking for bipolar symptoms past and present in the index patient, the presence of bipolar family history, or the presence of bipolar symptoms in family members. None of the studies systematically collected past records that often contain direct observations of mania or hypomania. The direct observation of signs and symptoms is the closest thing psychiatrists have to a "gold standard" for diagnosis. A clear example of this is the "Iowa 500" studies in which the diagnosis of Bipolar I was missed in one-third of patients until hospital records were reviewed (Winokur, Tsuang, & Crowe, 1982). The DSM-IV-TR specifically states "Information from other informants is often critical in establishing the diagnosis of Bipolar II Disorder." In one study, family members identified mania twice as frequently as the patients (47% vs. 22%; Ghaemi et al., 1999).

It is not unusual for patients to believe that hypomania is their "normal state" or "is when life is going their way." Clinically, it is not unusual for patients to present with prolonged depressed or mixed symptoms, making it difficult to find a clear hypomanic episode. It is not unusual for the patient's last clear hypomanic episode to have been 20 or more years ago. In one prospective study, only 40% of participants with Bipolar II experienced hypomanic episodes over a 10-year period (Coryell, Endicott, Maser, & Keller, 1995). The more remote the last hypomanic episode, the more difficult it is for patients to recall sufficient details to qualify for the diagnosis of hypomania.

Past records showing a response to treatment can indicate a correct diagnosis. A mere improvement of symptoms cannot be considered proof, but a complete long-term remission is certainly evidence that a diagnosis is correct. Long-term remission can also bring about diagnostic difficulties. Patients and families of patients who are euthymic and functional for long periods of time tend to question diagnosis, minimize or rationalize the significance of hypomanic symptoms, and or deny that they have bipolar disorder.

The SCID website discusses the importance of family informants, direct observations made over time, and review of past medical records.

4.2 | Limitations of the methods

4.2.1 | Reliability of Zimmerman's data analysis

In a "Head To Head" discussion with Dr. Zimmerman in the British Medical Journal, Smith and Ghaemi disagreed with Zimmerman's data analysis stating

> *In effect, bipolar disorder had been missed more than twice as frequently as it was mistakenly diagnosed (relative risk 2.23, 95% confidence interval 1.53 to 3.25). Unfortunately, low reliability in this study was reported as evidence of overdiagnosis when in fact the (validity) data suggest underdiagnoses. (Smith & Ghaemi, 2010)*

Despite being afforded the opportunity to respond to Smith and Ghaemi's criticism, Dr. Zimmerman only restated his previous position.

4.2.2 | A gold standard in the Bronze Age

Zimmerman and many others have asserted publicly and in writing that the SCID is the gold standard for diagnostic interviews (Kessler et al., 2004; Laino, 2008; Nordgaard, Revsbech, Sæbye, & Parnas, 2012). Is the SCID a true gold standard? The SCID website states "Unfortunately, a gold standard for psychiatric diagnoses remains elusive" (website). Gold standards are based on objective measurements that can be tested against an established standard. In psychiatry, there is no absolute established standard for diagnosis of bipolar disorder. Until biological markers are found, the validity of the SCID cannot be tested (Kessler et al., 2004).

The accuracy of structured interviews has been called into question both in general (Ventura, Liberman, Green, Shaner, & Mintz, 1998) and specifically with regard to the SCID. Earlier versions of the SCID missed a significant number of comorbid psychiatric conditions when substance dependence was present (Kranzler, Kadden, Babor, Tennen, & Rounsaville, 1996) that have not been retested in later SCID versions. Studies show "poor diagnostic performance" of the SCID when performed by nonpsychiatrists (Nordgaard et al., 2012). Of the four studies considered, only the Steward study used a psychiatrist to administer the SCID.

Multiple studies have found that structured interviews, including the SCID, have difficulty diagnosing Bipolar I (Benazzi, 2003b; Benazzi & Akiskal, 2003; Dunner, 1996; Dunner & Tay, 1993). Akiskal et al. showed that clinicians specifically trained to recognize Bipolar II were much better at finding the correct diagnosis than the SCID. Benazzi found that Bipolar II was missed ~50% of the time by the SCID when compared with a "semi-structured interview based on DSM-IV criteria conducted by an expert clinician" (Benazzi, 2003b). A 2009 study found significant "shortcomings" of structured clinical interviewers in general and the SCID specifically, finding them to have low reliability for detecting Bipolar II and low retest reliability (Miller, Johnson, & Eisner, 2009). The November 2011 update of the SCID formally acknowledged that the SCID was failing to adequately diagnose Bipolar II, stating "revisions have been made due to identified risks of false negatives." By extension, these shortcomings may have made the recognition of bipolar NOS even less robust.

Hirschfeld did not employ the full SCID, and the interviews were conducted over the telephone. There are no studies validating the reliability of the SCID performed by telephone. Nonverbal cues are often important in evaluating patients.

4.3 | Where are the cyclothymic patients?

According to the DSM-IV-TR, the lifetime prevalence of cyclothymic disorder is 0.4–1% and the lifetime prevalence of Bipolar I is 0.4–1.6% (APA & Association, 2000). This information would predict that the numbers of Bipolar I and cyclothymic patients should be similar; however, Hirschfeld failed to identify any patients with cyclothymic disorder. Zimmerman found only 2% of patients with cyclothymia compared with 29% of patients who were diagnosed with Bipolar I. The almost complete lack of cyclothymic diagnosis is concerning and reinforces the findings of other studies discussed that found weaknesses in the SCID.

4.3.1 | Family history as a validator

Zimmerman asserts that family history validates the findings of underdiagnoses. There are a number of potential problems with this assertion. First and foremost, the family history obtained from patients is secondhand (i.e., from the patient and not the family). Secondhand information is of questionable value. The second problem is a positive family history of bipolar disorder that has a specificity of 98% when predicting a diagnosis of bipolar disorder in an individual with depression or history of depression. However, the sensitivity is only 56%, and therefore the absence of a family history of bipolar disorder does not imply that an individual does not have bipolar disorder (Akiskal et al., 2000). Third, no criteria for evaluating the diagnosis of hypomania are found in the Family History Research Diagnostic Criteria. There was no discussion of how this lack of diagnostic criteria was handled. Fourth, and most important, the accuracy of the Family History Research Diagnostic Criteria has never been validated. In fact, its validity is called into question by the very Family History Research Diagnostic study cited by Zimmerman. Specifically, the Family History Research Diagnostic Criteria states that it yields good validity only when a relative is currently ill, which is "a situation that rarely occurs when family history data is obtained." The conclusion of the Family History Research Diagnostic Criteria study states "… underreporting remains a major problem of the family history method" (Andreasen, Endicott, Spitzer, & Winokur, 1977). The fifth problem is patients suffering from a bipolar disorder who "don't want to be bipolar," had high levels of denial, lacked insight, kept family secrets, and/or had other difficulties previously discussed would much more likely report a false negative family history for bipolar.

4.4 | Limitations of the boundaries of bipolar disorders

4.4.1 | The boundaries of bipolar disorders

Both the Goldberg and Stewart studies did not consider bipolar NOS or cyclothymia. However, they made no effort to limit their sample solely to patients previously diagnosed with Bipolar I or II. Many of the patients who reported a previous diagnosis of bipolar disorder

but failed to be diagnosed by the SCID may have been correctly diagnosed as having bipolar NOS or cyclothymia by their clinicians. Yet, under Goldberg's and Stewart's limited definition, these patients did not qualify for a bipolar diagnosis. This discrepancy alone might be grounds to seriously question the results of the Goldberg and Stewart studies.

In addition, Goldberg reported 63% of patients previously diagnosed as bipolar could not be similarly diagnosed because of "the identification of mood episodes that occurred solely within a 4-week period following significant substance misuse (i.e., mood disorders secondary to intoxication or withdrawal states, per criteria specified in DSM-IV)." This result appears to be an error. The DSM-IV has no such 4-week criterion. DSM-IV mania criteria E states that "The symptoms are not due to the direct physiological effects of a substance." For example, mood symptoms present during cocaine intoxication are not to be considered but, once past intoxication, manic symptoms can be considered as evidence of bipolar. A waiting period of 4 weeks is not part of the DSM-IV criteria and is not supported by the evidence base.

Hirschfeld found only 3% of patients were suffering from either bipolar NOS or cyclothymia. Zimmerman found a 26% rate. Both Hirschfeld's and Zimmerman's results fall short of the expected 1:1 ratio of the combined totals of Bipolar I and Bipolar II compared with the combined totals of bipolar NOS and cyclothymia (Miller, Dell'Osso, & Ketter, 2014), This shortfall highlights the concerns about the SCID.

4.4.2 | DSM–IV criteria versus evidence-based medicine

To practice evidence-based medicine, we must first practice evidence-based diagnosis. The DSM-IV criteria were published in January of 1994. The DSM-IV instructs us to incorporate new evidence. The DSM-IV and DSM-IV-TR both state that we should temper its use with new knowledge. The authors did not intend the criteria to be applied "statically" or "legalistically." All four studies used the DSM-IV 4-day requirement to diagnose hypomania (APA, 2013; APA & Association, 2000; APA & DSM-IV, 1994). The original selection of a minimum 4-day duration for hypomanic symptoms was arrived at by consensus in a committee rather than from empirical data (Parker, Graham, Synnott, & Anderson, 2014). Research shows that the average hypomanic episode lasts 2 days, with a 1–3-day range (Akiskal, 1996; Akiskal et al., 1979; Akiskal et al., 2000; Parker et al., 2014). There is an ongoing debate if the criteria for Bipolar II should be shortened from 4 to 2 days (Akiskal et al., 2000; Parker et al., 2014). Although this debate remains unresolved, there is sufficient evidence to consider a hypomania of 1–2 days as being sufficient for bipolar NOS. Zimmerman used a more conservative standard: "although we did not adopt a formal rule we diagnosed bipolar NOS in some patients with repeated hypomanic episodes that barely fell short of the duration requirements." Note that Zimmerman, without explanation, adds an additional hurdle for the diagnosis of bipolar NOS of "repeated hypomanic episodes." The DSM-IV does not require repeated episodes. Zimmerman acknowledged that a broader interpretation of bipolar NOS would have increased the diagnosis of bipolar disorders, yet failed to elucidate the number of patients who fell short of bipolar NOS or exactly how or why they fell short.

4.4.3 | Assertions without proof

Drs. Zimmerman and Goldberg have made assertions regarding why bipolar disorders are being overdiagnosed. Zimmerman asserts that "Clinicians are more inclined to diagnose conditions they feel more comfortable treating." He reasons because there are more medications to treat bipolar disorders than borderline personality disorders, clinicians are more inclined to diagnose bipolar disorder. He further posits that "this bias is reinforced by marketing messages from pharmaceutical companies." Further, he in essence states that clinicians are influenced by the many studies showing that it is underdiagnosed (Zimmerman et al., 2008). Dr. Zimmerman amplified this statement in the British Medical Journal stating "The marketing tactics of drug companies are absolutely a contributing factor" (Smith, Ghaemi, & Zimmerman, 2010). In 2010, Goldberg decried the "apparent lack of rigor with which practitioners formulate and diagnose complex mood disorders." Goldberg added "a multitude of disorders have become subject to erroneous reformulation as bipolar disorder and include a sizable wedge of the DSM-IV, including major depression, numerous anxiety disorders, etc." (Goldberg, 2010). It is concerning that these potentially volatile speculations were published in scientific journals. Many clinicians have found the statements to be unhelpful and believe they may have led to patients, therapists, other physicians, and even some psychiatrists to openly discount a diagnosis of bipolar disorder.

5 | CONCLUSIONS

Three studies concluded that bipolar disorders were being overdiagnosed. Hirschfeld did not conclude overdiagnosis of bipolar disorders in his study, but data from his study were used ad hoc by Dr. Zimmerman to support an assertion of overdiagnosis. Studies that conclude bipolar disorders are overdiagnosed are retrospective and call for further studies to validate their findings. There have been no further studies. No branch of medicine is free from misdiagnosis. The term overdiagnosis in these studies is used as a substitute for an unacceptable high rate of false positives, yet these studies fail to adequately establish the overdiagnosis of bipolar disorders.

All these studies arrived at a diagnosis after a single interview without a systematic family history or past records. The DSM-IV-TR specifically warns that one interview is inadequate to rule out bipolar disorders and emphasizes the importance of obtaining a family history. Collecting past records is crucial in a disorder in which lack of insight, state dependent memory, remoteness of hypomanic episodes, lack of a common vocabulary, denial, shame, transference issues, and wanting to please the interviewer can significantly interfere with the diagnostic process.

The Family History Method used by Zimmerman is an instrument that is self-admitted to be of limited value. The original study of the Family History Method openly acknowledges that underreporting is a major problem. To further complicate this issue and add to the inadequacy of this method, the instrument was developed prior to the formal recognition of Bipolar II and has no specific criteria for hypomania. The Family History Method may not be the best choice to use as a validator.

The assertion that the SCID represents a gold standard is unsupportable. Until accurate biological markers are developed, the assertion that the SCID is a gold standard is not even testable. The makers of the SCID openly acknowledge that the SCID is not a gold standard. Numerous studies published well before the overdiagnosis studies were conducted show the SCID to be weak for diagnosing Bipolar II. None of these studies were adequately discussed in any of the overdiagnosis studies.

Any one of the aforementioned factors may be sufficient to question the conclusions of these studies. Taken together, these factors make the assertion that the bipolar disorders are being overdiagnosed less than certain.

5.1 | There is more

Zimmerman's, Stewart's, and Goldberg's conservative boundaries are at odds with both the evidence-base and DSM-IV criteria. The boundaries used in these studies of what constitutes a diagnosis of bipolar disorder may have been more conservative than those of the clinicians. Indeed, Zimmerman acknowledges this possibility. Both the Goldberg's and Zimmerman's studies applied criteria that were inconsistent with the DSM-IV. Goldberg's and Stewart's failure to consider bipolar NOS or cyclothymia as valid diagnoses is particularly curious. A narrow definition of a disease compared with a broader definition is always going to result in fewer diagnoses.

Ultimately, the studies that assert that bipolar disorders are being overdiagnosed may be rejected on a more fundamental basis. These studies are comparison studies. They compare the diagnosis of bipolar disorder made by clinicians to the diagnosis made by the SCID. All of the studies assumed that the SCID diagnosis was right and the clinician's diagnosis was wrong despite the possibility that clinicians may have had the advantage of direct observations made over time, more direct family histories, and past records. The four studies collected virtually no information of how or why the clinician arrived at the diagnosis of bipolar disorder. Indeed, because these studies were largely based on patients' self-reports, we cannot even be sure if the patients were previously diagnosed with a bipolar disorder. Ultimately, all four studies used to assert that bipolar disorders are being overdiagnosed compare a single SCID interview to a clinician's diagnosis. Yet, the clinician side of the comparison is virtually without data, making the comparison absurd.

Over time, Drs. Zimmerman and Goldberg have abandoned their scientific neutrality to assert with certainty that bipolar disorder is being overdiagnosed. Dr. Zimmerman has done so despite the challenge to his data interpretation that he has left unanswered (Smith et al., 2010) and an update of the SCID that specifically admits to an unacceptable false negative rate of the diagnoses of Bipolar II. Their opinions with regard to drug companies and the skills of clinicians are beyond the pale. These assertions, offered without evidence, may have caused a loss of confidence in psychiatry, psychiatrists, and, in particular, rank-and-file clinicians. Clinically, because these studies were published, it is not unusual to find psychologists, master's level therapists, and nonpsychiatric physicians who openly disagree with a diagnosis of bipolar disorder. Even worse, these professionals may covertly disagree with the diagnosis, thereby undermining the psychiatrist to the detriment of patients and their families. The opinions of these authors may be the result of an abreaction to the rapid growth of the numbers of patients being recognized with a bipolar disorder.

The authors acknowledge some of the concerns discussed above but have understated their significance. Many other concerns seem to have gone unrecognized; for example, the conservative boundaries used to diagnosis bipolar disorder, the lack of validity of the Family History Method, and the lack of data to make adequate comparison.

The boundaries of bipolar disorder are not expanding. Rather, our understanding of bipolar disorders is deepening. It appears that we are coming closer to Kraepelin's original definition. We all desire a clear understanding of the boundaries of bipolar disorders. We are finding, just as Kraepelin did over 120 years ago, that bipolar disorders may "include the slight and slightest colorings of mood" and may "pass over without sharp boundaries to personal predisposition" (Kraepelin, 1921/1898). Until the time we develop biological markers for bipolar disorders, we are going to have to live with the ambiguity that nature has presented to us.

Understanding the newest research and statistical methods is critically important in psychiatry because the present state of knowledge compels us to rely on potentially subjective information when assessing patients. Concurrently, we should not lose sight of the importance of evaluating the more fundamental aspects of research.

FUNDING

Self-funded.

DECLARATION OF INTEREST STATEMENT

There are no conflicts of interest, the author is on the speaker bureau for Allergan.

There are no prior presentations of this study. The author alone is responsible for the content and writing of the article.

ACKNOWLEDGEMENTS

The author thank the following individuals for assisting with ideas, editing, and suggestions: Daniel Lieberman, MD, Professor of Psychiatry and Behavioral Sciences, Department of Psychiatry and Behavioral Sciences, George Washington University, Washington DC; Rabbi Lawrence Denmark, Ed. D, The Depression & Bipolar Clinic of Colorado, Fort Collins, CO; James Phelps, MD, Director, http://psycheducation.org, Staff Psychiatrist, Samaritan Mental Health, Corvallis, OR, and Bipolar Section Editor, *Psychiatric Times*; Christopher Aiken, MD, Director, Mood Treatment Center, Winston-Salem, NC, and Instructor in Clinical Psychiatry, Wake Forest University School of Medicine, Winston-Salem NC.

ORCID

Tammas Kelly http://orcid.org/0000-0001-8805-975X

REFERENCES

Akiskal, H. S. (1996). The prevalent clinical spectrum of bipolar disorders: Beyond DSM-IV. *Journal of Clinical Psychopharmacology*, *16*(2), 4S–14S.

Akiskal, H. S., Bourgeois, M. L., Angst, J., Post, R., Möller, H.-J., & Hirschfeld, R. (2000). Re-evaluating the prevalence of and diagnostic composition within the broad clinical spectrum of bipolar disorders. *Journal of Affective Disorders*, *59*, S5–S30.

Akiskal, H. S., Maser, J. D., Zeller, P. J., Endicott, J., Coryell, W., Keller, M., … Goodwin, F. (1995). Switching from 'unipolar' to bipolar II: An 11-year prospective study of clinical and temperamental predictors in 559 patients. *Archives of General Psychiatry*, *52*(2), 114–123.

Akiskal, H. S., Rosenthal, R. H., Rosenthal, T. L., Kashgarian, M., Khani, M. K., & Puzantian, V. R. (1979). Differentiation of primary affective illness from situational, symptomatic, and secondary depressions. *Archives of General Psychiatry*, *36*(6), 635–643.

Akiskal, H. S., Walker, P., Puzantian, V. R., King, D., Rosenthal, T. L., & Dranon, M. (1983). Bipolar outcome in the course of depressive illness: Phenomenologic, familial, and pharmacologic predictors. *Journal of Affective Disorders*, *5*(2), 115–128.

Albanese, M. J., Clodfelter, R. E. Y. N. O. L. D. S. C. Jr., Pardo, T. B., & Ghaemi, S. N. (2006). Underdiagnosis of bipolar disorder in men with substance use disorder. *Journal of Psychiatric Practice®*, *12*(2), 124–127.

Andreasen, N. C., Endicott, J., Spitzer, R. L., & Winokur, G. (1977). The family history method using diagnostic criteria: Reliability and validity. *Archives of General Psychiatry*, *34*(10), 1229–1235.

APA (2013). *DSM-5 diagnostic and statistical manual of mental disorders (DSM-5®)*. Washington, DC: American Psychiatric Publishing.

APA, & Association, American Psychiatric (2000). *DSM-IV-TR* (Vol. 75). Washington, DC: American Psychiatric Association.

APA, & DSM-IV, American Psychiatric Association. Task Force on (1994). *DSM-IV*. Washington, DC: American Psychiatric Association.

Benazzi, F. (2003a). Frequency of bipolar spectrum in 111 private practice depression outpatients. *European Archives of Psychiatry and Clinical Neuroscience*, *253*(4), 203–208. https://doi.org/10.1007/s00406-003-0433-6

Benazzi, F. (2003b). Diagnosis of bipolar II disorder: A comparison of structured versus semistructured interviews. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *27*(6), 985–991.

Benazzi, F., & Akiskal, H. S. (2003). Refining the evaluation of bipolar II: Beyond the strict SCID-CV guidelines for hypomania. *Journal of Affective Disorders*, *73*(1), 33–38.

Braw, Y., Sitman, R., Sela, T., Erez, G., Bloch, Y., & Levkovitz, Y. (2012). Comparison of insight among schizophrenia and bipolar disorder patients in remission of affective and positive symptoms: Analysis and critique. *European Psychiatry*, *27*(8), 612–618.

Coryell, W., Endicott, J., Maser, J. D., & Keller, M. B. (1995). Long-term stability of polarity distinctions in the affective disorders. *The American Journal of Psychiatry*, *152*(3), 385–390.

Dunner, D. L., Gershon, E. S., & Goodwin, F. K. (1976). Heritable factors in the severity of affective illness. *Biological Psychiatry*, *11*(1), 31–42.

Dunner, D. L. (1996). In T. A. Widiger, A. J. Frances, H. A. Pincus, R. Ross, M. B. First, & W. W. Davis (Eds.), *Bipolar depression with hypomania (Bipolar II) DSM-IV Sourcebook* (Vol. 2) (pp. 53–63). Washington, DC: American Psychiatric Association Press.

Dunner, D. L., & Tay, L. K. (1993). Diagnostic reliability of the history of hypomania in bipolar II patients and patients with major depression. *Comprehensive Psychiatry*, *34*(5), 303–307. https://doi.org/10.1016/0010-440X(93)90015-V

Ghaemi, S. N., Boiman, E. E., & Goodwin, F. K. (2000). Diagnosing bipolar disorder and the effect of antidepressants: A naturalistic study. *J Clin I'svchiatrv*, *61*(10).

Ghaemi, S. N., Ko, J. Y., & Goodwin, F. K. (2002). "Cade's disease" and beyond: Misdiagnosis, antidepressant use, and a proposed definition for bipolar spectrum disorder. *The Canadian Journal of Psychiatry*, *47*(2), 125–134.

Ghaemi, S. N., Sachs, G. S., Chiou, A. M., Pandurangi, A. K., & Goodwin, F. K. (1999). Is bipolar disorder still underdiagnosed? Are antidepressants overutilized? *Journal of Affective Disorders*, *52*(1), 135–144.

Ghouse, A. A., Sanches, M., Zunta-Soares, G., Swann, A. C., & Soares, J. C. (2013). Overdiagnosis of bipolar disorder: A critical analysis of the literature. *The Scientific World Journal*, *2013*, 1–5.

Goldberg, J. F. (2010). Ruling in and ruling out 'caseness' in the bipolar spectrum: Implications for initial screening. *Bipolar Disorders*, *12*(5), 539–540.

Goldberg, J. F. (2015). Mixed depression: A farewell to differential diagnosis? *The Journal of Clinical Psychiatry*, *76*(3), 378–380.

Goldberg, J. F., Garno, J. L., Callahan, A. M., Kearns, D. L., Kerner, B., & Ackerman, S. H. (2008). Overdiagnosis of bipolar disorder among substance use disorder inpatients with mood instability. *The Journal of Clinical Psychiatry*, *69*(11), 1751–1757.

Hirschfeld, R. M., Calabrese, J. R., Weissman, M. M., Reed, M., Davies, M. A., Frye, M. A., … McNulty, J. P. (2003). Screening for bipolar disorder in the community. *The Journal of Clinical Psychiatry*, *64*(1), 53–59.

Hirschfeld, R., Lewis, L., & Vornik, L. A. (2003). Perceptions and impact of bipolar disorder: How far have we really come? Results of the national depressive and manic-depressive association 2000 survey of individuals with bipolar disorder. *The Journal of clinical psychiatry*, *64*, 161–174.

Hirschfeld, R. M. A., Cass, A. R., Holt, D. C. L., & Carlson, C. A. (2005). Screening for bipolar disorder in patients treated for depression in a family medicine clinic. *The Journal of the American Board of Family Practice*, *18*(4), 233–239.

Kessler, R. C., Abelson, J., Demler, O., Escobar, J. I., Gibbon, M., Guyer, M. E., … Walters, E. E. (2004). Clinical calibration of DSM-IV diagnoses in the World Mental Health (WMH) version of the World Health Organization (WHO) Composite International Diagnostic Interview (WMH-CIDI). *International Journal of Methods in Psychiatric Research*, *13*(2), 122–139.

Kraepelin, Emil. (1921). *Manic-depressive insanity and paranoia*. Edinburgh: E & S Livingstone (Original work published 1898).

Kranzler, H. R., Kadden, R. M., Babor, T. F., Tennen, H., & Rounsaville, B. J. (1996). Validity of the SCID in substance abuse patients. *Addiction*, *91*(6), 859–868.

Laino, C. (2008). Bipolar disorder often misdiagnosed: Presented at APA. Retrieved from http://www.peervoice.com/bipolar-disorder-often-misdiagnosed-presented-apa

Látalová, K. (2012). Insight in bipolar disorder. *Psychiatric Quarterly*, *83*(3), 293–310.

Miller, C. J., Johnson, S. L., & Eisner, L. (2009). Assessment tools for adult bipolar disorder. *Clinical Psychology: Science and Practice*, *16*(2), 188–201.

Miller, S., Dell'Osso, B., & Ketter, T. A. (2014). The prevalence and burden of bipolar depression. *Journal of Affective Disorders*, *169*, S3–S11.

Mitchell, P. B. (2012). Bipolar disorder: The shift to overdiagnosis. *The Canadian Journal of Psychiatry*, *57*(11), 659–665.

Mitchell, P. B., Goodwin, G. M., Johnson, G. F., & Hirschfeld, R. (2008). Diagnostic guidelines for bipolar depression: A probabilistic approach. *Bipolar Disorders*, *10*(1p2), 144–152.

Nordgaard, J., Revsbech, R., Sæbye, D., & Parnas, J. (2012). Assessing the diagnostic validity of a structured psychiatric interview in a first-admission hospital sample. *World Psychiatry*, *11*(3), 181–185.

Parker, G., Graham, R., Synnott, H., & Anderson, J. (2014). Is the DSM-5 duration criterion valid for the definition of hypomania? *Journal of Affective Disorders*, *156*, 87–91.

Phelps, J., Angst, J., Katzow, J., & Sadler, J. (2008). Validity and utility of bipolar spectrum models. *Bipolar Disorders*, *10*(1p2), 179–193.

Smith, D., Ghaemi, N., & Zimmerman, M. (2010). Is underdiagnosis the main pitfall when diagnosing bipolar disorder? *British Medical Journal*, *340*(7748), 686–687.

Smith, D. J., & Ghaemi, N. (2010). Is underdiagnosis the main pitfall when diagnosing bipolar disorder? Yes. *British Medical Journal*, *340*, c854. https://doi.org/10.1136/bmj.c854

Smith, D. J., Griffiths, E., Kelly, M., Hood, K., Craddock, N., & Simpson, S. A. (2011). Unrecognised bipolar disorder in primary care patients with depression. *The British Journal of Psychiatry*, *199*(1), 49–56.

Stewart, C., & El-Mallakh, R. S. (2007). Is bipolar disorder overdiagnosed among patients with substance abuse? *Bipolar Disorders*, *9*(6), 646–648.

Ventura, J., Liberman, R. P., Green, M. F., Shaner, A., & Mintz, J. (1998). Training and quality assurance with the structured clinical interview for DSM-IV (SCID-I/P). *Psychiatry Research*, *79*(2), 163–173. https://doi.org/10.1016/S0165-1781(98)00038-9

Website, SCID official. What is the "validity" of the SCID-I? 2011. Retrieved from http://www.scid4.org/psychometric/scidI_validity.html

Winokur, G., Tsuang, M. T., & Crowe, R. R. (1982). The Iowa 500: Affective disorder in relatives of manic and depressed patients. *The American Journal of Psychiatry*. https://doi.org/10.1176/ajp.139.2.209

Zimmerman, M. (2010). Problems diagnosing bipolar disorder in clinical practice. *Expert Review of Neurotherapeutics*, *10*(7), 1019–1021.

Zimmerman, M., Ruggero, C. J., Chelminski, I., & Young, D. (2008). Is bipolar disorder overdiagnosed? *The Journal of Clinical Psychiatry*, *69*(6), 935–940. Web of Science

How to cite this article: Kelly T. Prospective: Is bipolar disorder being overdiagnosed? *Int J Methods Psychiatr Res.* 2018;27:e1725. https://doi.org/10.1002/mpr.1725